Communication Networks

Generalized Session Models for Wireless Cellular Networks

Harro L. Hartmann

Institute for Computer and Communication Network Engineering

D-38106 Braunschweig, Germany

h.hartmann@tu-bs.de

Abstract - This contribution deploys generalized basic relations of session models, which remain completely independent of the possible underlying technologies. First, a renewal session trajectory model is proposed and yields a mean value theorem. This confirms that the mean number of handoffs remains insensitive with respect to the density distributions of the participating residence and session times with vanishing forced terminations. In contrast to rather complicated and pure transform domain solutions of the past a hybrid original-transform-domain approach is proposed. This keeps the relations physically transparent and facilitates the consideration of handoff blockings and forced terminations. Second, a further theorem based on an appropriate inversion of the transform domain densities shows that the state probabilities of handoffs may be expressed in an explicit symbolic form if generic Gamma distributed session and residence time durations are assumed. Third, an estimated mean Diameter protocol rate for generally distributed session and residence times including forced terminations proves to be given in an explicit form too. Finally, keying moments and complementary distribution functions of generalized handoff process are symbolically derived and enumerated.

Keywords - Wireless Cellular Networks, Handoff Statistics in PCS, Residence Capacity of Sessions, Session Time Changes by forced Terminations.

1. Introduction

Mobility, performance and accounting are key issues for the expanding wireless cellular networks. Mobility implies that a cellular call between voice terminals or a session with respect to computer oriented transfers, originating from anywhere, any time within the service area, would be able to maintain the same call or session without service interruption, while in motion. The performance concerns the switching capability (e.g. channel assignment, handover, and power control) of the Base Station Controller (BSC) for all connected cells including GSM Base Transceiver Stations or UMTS Nodes B respectively. Here the obtainable Erlangian traffic capacity with respect to required GoS and QoS requirements results from the available channel group capacity. Authentication, Authorization and Accounting (AAA) form related process components and are performed by so-called AAA systems operating on the Radius or Diameter Protocols respectively. Here, signaling rates, call and handover blockings remain important issues.

Since about two decades a huge amount of contributions has been devoted to refined network models [1- 6]. Zonoozi and Dassanayake considered in [1] cell residence times of new and handover calls. They fitted the random movement of mobile stations by a generalized Gamma distribution and validated that independent from the user's call holding time the cell holding time becomes exponentially distributed. This result is remarkably because it is keying for the obtainable traffic capacity of each cell switching capability within a provisioning cluster. Instead of the resulting G/M/n system Hong and Rappaport proposed the classical M/G/n cell traffic and performance model implying different priority schemes in [2]. The key design parameters are therefore the forced termination probabilities for handoff calls and blocking probabilities for new calls. In essence, one prefers to make already processed handoff calls more successfully than new calls requiring service. Simulation studies of different dynamic prioritization schemes, using real-time mobile positioning information, were presented by Soh and Kim in [3].

This contribution will assume both types of blocking as required capacity assignment parameters (GoS) and develops a renewal process model of the resulting handoff statistic within provisioning cell clusters subject to generic session and residence time interrelations. Thus, this approach focusses on counting of events which shortcuts a huge amount of analytical details. But in contrast to pure renewals we are faced with a delayed and defective renewal process because residual residence time delays occur during the session set up and handoffs may be blocked before the session time expires or if the mobile leaves the cell cluster. Within these constraints we may define a counting process and thus preserve the renewal orientation. A consequence of this approach is that a closed and intuitive transparent solution can be provided even for generic Gamma function distributed process components.

In the past many valuable papers proceeded to more or less generally distributed total holding and residence times. Considerable investigations in this area subject to blocking and forced terminations of calls were made by Fang and Clamtac in [4 - 5]. In these and further contributions [6 - 7]. the authors preferred an appealing unified transform domain approach. But unfortunately, this method requires expensive analytical complex plane refinements. Either transcendent Laplace transforms occur or extensive fittings of realistic field date are desired. Thus, most desirable results remain numerically expensive or strongly depend on a large set of parameter combinations respectively. Of course the physical transparency of the system response remains hidden. For example, multiple partial derivatives performed by available computer algebra systems may result in pages of expressions and uncountable simplification alternatives. But, the more difficult the distributions of session and residence times appear, the more simply the solution steps should be. So the approach here differs from [4 - 7] by the dominant original-domain solution method which admits irrational and transcendent functions at an considerably reduced amount of space. Suggestions and first applications with respect to AAA were proposed in [8 - 9].

This contribution is organized as follows. Chapter 2 establishes a novel unified renewal model and a mean value theorem for generally distributed process interrelations. Chapter 3 covers basic handoff probability distributions and selected moments which justify a second theorem. In Chapter 4 generic generalizations of handoffs statistics are considered. Selected applications with respect to Radius and Diameter protocol subject to forced terminations are given in Chapter 5. The conclusions highlight possible limitations versus achieved advantages and recommendations for further work. Finally, the Appendix explains selected basic relations and the mean effective session time with forced terminations accompanied by a third theorem.

2. The Renewal Process oriented Model and Mean Number of Handoffs in a Session

We assume that the ensemble of mobile terminals (MTs) generates Poisson session arrivals. With mobile terminals, we accompany an arrived session via a random number of visited cells within a cluster subject to limited residence until the tagged session terminates or the moving terminal leaves the provisioning area. If the residence within a selected cell exhausts a handoff appears and a session handover to another cell occurs. Since each session starts at an arbitrary point within the first visited cell and an accepted session doesn't terminate within this cell the first handoff occurs after a residual residence time $\tilde{R}$. Then the session time S and further residence times R are assumed to be independent identically distributed (iid) random variables. Now, each accepted session forms a trajectory as seen by **Fig. 1**, similar to that in [1] but more dedicated to an intended session model.

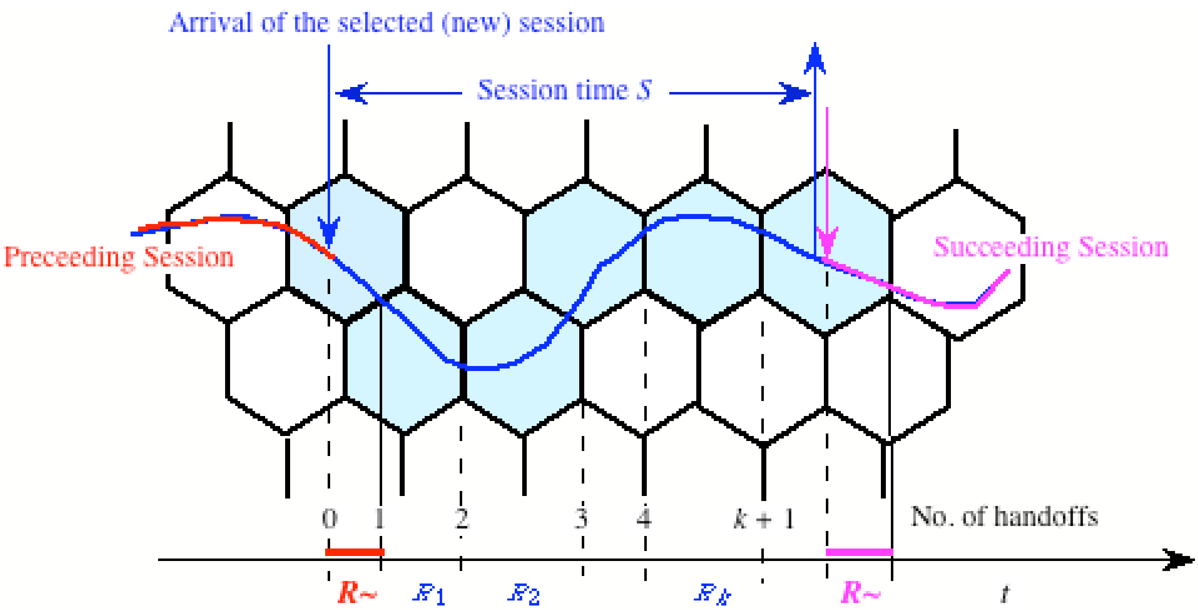

Fig. 1. The MT space-time trajectory within a cell cluster and residence times capacity of a selected session, R~: New session residual cell residence time, R: Handoff cell residence time.

In more detail we distinguish one residual residence time and k complete residence times including $(k+1)$ handoffs for every accepted new session. Any possible handover blocking is considered later. Our objective here is the statistical estimation of the complete number of handoffs K. After the delay by a renewal oriented residual time $\tilde{R}$ the possible following inter-occurrence times R may be described by a pure renewal process. Then $\tilde{R}+R$ form a delayed renewal process where the distribution of $\tilde{R}$ differs from that of R in general but both events remain independently distributed.

The huge amount of possible process combinations requires effective communication. Therefore we remember D. G. Kendall's system characterization and proceed to an extended application. On the one hand we observe the ensemble of session trajectories or mobile paths. For example we consider a cluster of cells each of which performs like a standard M/G/n system. But on the other side we focus our attention on selected trajectories per virtual channel only. So we introduce the non-standard designation, namely "Session/Residence/Cluster-capacity" abbreviated by "S/R/C" for the underlying cellular system. Observe that the first both entries characterize the interfering processes and the last entry forms the possible constraint. If not expressly denoted we assume C to be infinitely subject to possible truncations considered in Chapter 4.

By these assumptions we get exclusively

$$T_k = \tilde{R} + \sum_{j=1}^{k} R_j \text{ with probability } P\{K=k\},\ k=1,2,3,..\infty \text{ and} \tag{1a}$$

$$T_0 = \tilde{R} \text{ with probability } P\{K=0\} = p_0 = 1-\sum_{k=1}^{\infty} p_k = 1-P\{K>0\}. \tag{1b}$$

Now, these equations define a delayed renewal process with discrete epochs T_k and the number of renewals (residences) K in $(0,t]$. In essence, the sequence $\{R_j\}$ forms an ordinary renewal or recurrent process, which also may be represented by a counting process $\{K(t), t \geq T_0\}$ generated by the inter-occurrence times R_j. In essence, $K(t)$ is the largest integer $k>0$ for which $T_j \leq t$ and represents the number of events up to time t. Thus, the events $\{K(t) \geq j\}$ and $\{T_j \leq t\}$ are equivalent; hence and by $\{K \geq j+1\}$ equivalent to $\{K > j\}$ we obtain $P\{K(t)>k\} = P\{T_k < t\}$. Now, we generalize these equivalences by the assumption that t appears to be the random variable S and obtain including the multiple convolution $g_R(x,k) = f_{\tilde{R}}(x) * f_R(x)^{(*k)}$

$$P\{S>T_k\} = P\{K(S)>k\} = \int_0^{\infty} \bar{F}_S(x) \cdot g_R(x,k)\,dx, \qquad \text{(G/G/C)} \tag{2a}$$

completed by the zero state or delay probability

$$P\{S \leq T_0) = P\{K=0\} = 1-P\{K>0\}. \tag{2b}$$

Note that any distribution function of the counting process $K(S)$ and similarly that of T_k focus on the renewals and the renunciation of the complete time domain simplifies advanced modeling considerably.

Let $f_X^*(s) = L\{f_X(x)\}$ denote the Laplace Transform of $f_X(x)$ and $L^{-1}\{f_S^*(s)\} = f_X(x)$ its inverse. Then, $g_R(x,k) = L^{-1}\left\{f_{\tilde{R}}^*(s)[f_R^*(s)]^k\right\}$. This inverse transform of the delayed renewal process exists for many important reference applications, see Chapter 4. Now, the complete form of (10b) is

$$P\{K \geq k+1\} \equiv P\{K > k\} = \int_0^\infty \bar{F}_S(x) \cdot L^{-1}\left\{\frac{1 - f_R^*(s)}{m_{1R}s}[f_R^*(s)]^k\right\}dx,\ k \geq 0\,. \tag{3}$$

Note the keying elimination of $f_{\tilde{R}}^*(s)$ by $f_R^*(s)$. The focus on the statistic of K as a hybrid transform-original domain solution of the delayed renewal process remains much more transparent and compact than many other complete transform-domain approach for different handoff probabilities as proposed in [4 - 6]. We will experience this below up to cases of irrational and transcendent Laplace transforms of the participating processes. Furthermore we expect that the incorporation of new call and handover call blockings simplify the resulting extended solutions too, cf. Appendix A2 and A3. First, instead of solving for $P\{K > k\}$ we may directly proceed by (3) to

$$m_{1K} = \sum_{k=0}^{\infty} \bar{F}_K(k) = \int_0^\infty \bar{F}_S(x) \cdot L^{-1}\left\{\frac{1}{m_{1R}s}\sum_{k=0}^{\infty}[1 - f_R^*(s)][f_R^*(s)]^k\right\}dx \text{ and}$$

$$m_{1K} = \frac{1}{m_{1R}}\int_0^\infty \bar{F}_S(x) \cdot L^{-1}\left\{\frac{1}{s}\right\}dx = \frac{1}{m_{1R}}\int_0^\infty \bar{F}_S(x) \cdot U(x)dx \equiv \frac{m_{1S}}{m_{1R}} = \frac{\mu_R}{\mu_S}\ \text{, (G/G/C)} \tag{4a}$$

where $U(x)$ is the unit step function. This proves the following *mean value Theorem*.

Theorem I: The mean number of handoffs caused by K independently general distributed residence times subject to general distributed session times depends only on the ratio of the first moments and not on the form of the involved distribution functions.

Comments: The most generally valid moment ratio $(m_{1S} / m_{1R}) = \mu_R / \mu_S$ is often used as a mobility rate measure and may be allocated to service specific sources and links of open networks. Observe that the probability distribution of K may vary within the same mean for arbitrary session and residence time distribution functions (DFs). Moreover, if the entire provisioning area has C cells, the total non-blocked average new session attempt rate can be given by $\lambda = C\lambda_{Cell}$. Then the total average handoff session rate amounts to $\lambda_{ho} = m_{1K}C\lambda_{Cell}$. Assuming that these traffic components are equally distributed among all cells, the ratio of the average carried handoff attempts to the average new call originating rate becomes

$(\lambda_{ho} / \lambda) \equiv m_{1K}$. Thus, the ratio of average total traffic λ_t to the new session attempt rate amounts to

$$\lambda_t = \lambda + \lambda_{ho} = (1 + m_{1K})\lambda . \tag{4b}$$

3. Basic Relations for the Probability Distribution of *K*

According to (2) the state probabilities are defined by

$$p_k = P\{K > k-1\} - P\{K > k\} = \bar{F}_K(k-1) - \bar{F}_K(k) \,,\ k = 1,2,...\infty \text{ but}$$

$$p_0 = P\{K = 0\} = 1 - P\{K > 0\} \,,\ p_0 + \sum_{k=1}^{\infty} p_k \equiv 1 .$$

Here, the most general **G/G/C** cases require the evaluation of

$$p_k = \int_0^{\infty} \bar{F}_S(x) \cdot L^{-1} \left\{ \frac{1 - f_R^*(s)}{m_{1R} s} \{[f_R^*(s)]^{k-1} - [f_R^*(s)]^k\} \right\} dx \text{ and} \tag{5a}$$

$$p_0 = 1 - P\{K > 0\} = 1 - \int_0^{\infty} \bar{F}_S(x) \cdot L^{-1} \left\{ \frac{1 - f_R^*(s)}{m_{1R} s} \right\} dx \text{ respectively.} \tag{5b}$$

Unfortunately, the distributions of K depend on the involved processes S and R and a transparently simple G/G/C approach requires modular decompositions. Therefore we attack the problem stepwise. Fortunately, favorable basic scenarios can be established by M/G/C which forms an important reference for three reasons. First, it provides physically transparent checks down to the M/M/C base case. Second, important cases of hyper-exponential distributions are covered, [10]. Third, the M/G/C scenario forms a base case for modular extensions to G/G/C scenarios which are considered in Chapter 4.

Now, for many processes R we may assume that their Laplace transformation $f_R^*(s) = L\{f_R(x)\}$ exists or may at least be approximated. Here $\bar{F}_S(x) = e^{-\mu_S x}$ and the integral $\int_0^{\infty} e^{-\mu x} g(x) dx \equiv L\{g(x)\}\big|_{s=\mu} = g^*(\mu)$ defines a single value of $g^*(s)$ and solves (5). So we can introduce $p = f_R^*(\mu_S)$. Then and in agreement to (4)

$$p_k = \frac{\mu_R}{\mu_S}(1-p)(p^{k-1} - p^k) \,,\ k = 1,2,...\infty \,, p_0 = 1 - \frac{\mu_R}{\mu_S}(1-p) \text{ and} \tag{6a}$$

$$P\{K > k\} = \sum_{j=k+1}^{\infty} p_j = \frac{\mu_R}{\mu_S}(1-p)p^k \,,\ p = f_R^*(\mu_S) . \tag{6b}$$

These equations establish the M/G/C probability distribution and justify the following

Theorem II: If the Laplace transform (not necessary rational) of the iid residence times exists, then the M/G/C – handoff CDF subject to $m_{1K} = \mu_R / \mu_S$ and a remaining specification of the generally defined components $p \equiv f_R^*(\mu_S) < 1$ is defined by $P\{K > k\} \equiv m_{1K}(1-p)p^k$.

Comments: The simple constraint m_{1K} saves the introduction of generally distributed residence times. Furthermore, the higher order M/G/C moments are accessibly by (6) and yield

$$m_{\ell K} = \frac{\mu_R}{\mu_S} \sum_{k=1}^{\infty} k^{\ell} (1-p)(p^{k-1} - p^k) \equiv \sum_{k=0}^{\infty} [(k+1)^{\ell} - k^{\ell}] P\{K > k\} . \tag{7a}$$

For the first two moments we prefer the non expanded sum form (7a). Here we maintained the commonly used complementary distribution function (CDF) substitution $\bar{F}_K(k) = P\{K > k\}$ instead of $P\{K \geq k-1\}$ subject to $k \geq 1$. For the first two moments we have

$$m_{1K} \equiv \sum_{k=0}^{\infty} P\{K > k\} = \frac{\mu_R}{\mu_S} \text{ again, } m_{2K} \equiv \sum_{k=0}^{\infty} (2k+1) P\{K > k\} = \frac{\mu_R}{\mu_S} \frac{1+p}{1-p} . \tag{7b}$$

Thus, the squared coefficient of variation yields

$$c_K^2 = \frac{m_{2K}}{m_{1K}^2} - 1 = \frac{\mu_S}{\mu_R} \frac{1+p}{1-p} - 1 = \frac{1}{m_{1K}} \frac{1+p}{1-p} - 1 . \tag{7c}$$

Note that in the M/M/C case both equations model the ordinary geometric distribution (GEO). Here, the event probability is determined by $p = f_R^*(\mu_S) = [\mu_R / (\mu_S + \mu_R)]$ or $(\mu_R / \mu_S) = p / (1-p)$. Then

$$f_R(x) = \mu_R e^{-\mu_R x} \equiv f_{\tilde{R}}(x),\ m_{1R} = 1/\mu_R,\ c_R = 1 .$$

The handoff probability distribution and its CDF become respectively

$$p_k = (1-p)p^k , \qquad P\{K > k\} = \sum_{j=k+1}^{\infty} (1-p)p^j = p^{k+1} . \quad \text{(M/M/C)} \tag{8a}$$

For the moment dependencies on p we have

$$m_{1K} = \frac{\mu_R}{\mu_S} \equiv \frac{p}{1-p} = m_{1\text{GEO}} \text{ and } c_K = \left(\frac{1+p}{p} - 1 \right)^{1/2} \equiv \frac{1}{\sqrt{p}} = c_{\text{GEO}} . \tag{8b}$$

Guided by (6) and (7) we can easily proceed to useful generalizations towards M/G/C systems now. First, we consider M/Γ/C systems assuming Gamma (Γ) distributed residence times. Here the distribution density implies the Euler Gamma function $\Gamma(\gamma) = \int_0^{\infty} t^{\gamma-1} e^{-t} dt$, the shape factor $\gamma > 0$ and the scale parameter $\alpha > 0$ according to

$$f_R(x) = \frac{\alpha^{\gamma} x^{\gamma-1} e^{-\alpha x}}{\Gamma(\gamma)},\ m_{1R} = \frac{\gamma}{\alpha} = \frac{1}{\mu_R},\ c_R = \frac{1}{\sqrt{\gamma}}, \text{ and } p = f_R^*(\mu_S) = \left(\frac{\alpha}{\mu_S + \alpha} \right)^{\gamma} . \tag{9}$$

This density reduces to exponentially distributed residence times for $\gamma = 1$ and $\alpha = \mu_R$ again which may be used for comparisons. Substitution of $\alpha = \mu_R \gamma$ reduces the moments to the shape factor versus the mobility ratio μ_R / μ_S only. Thus, we obtain for the underlying model $(\mu_R / \mu_S) = \kappa$ again but

$$c_K = \left[\frac{1}{\kappa} \frac{1 + [\kappa / (\kappa + c_R^2)]^{1/c_R^2}}{1 - [\kappa / (\kappa + c_R^2)]^{1/c_R^2}} - 1 \right]^{\frac{1}{2}}. \quad (\mathrm{M} / \Gamma / \mathrm{C}) \tag{10}$$

Now, **Fig. 2** shows the coefficient of variation c_K for different values of c_R, versus the normalized mobility rate κ. Observe firstly that for sufficient large κ the session time coefficient of variation $c_K \to c_S = 1$. Second, smaller mobility rates cause significant $c_K > c_S$ caused by c_R, thus different residence time densities cause different c_K too.

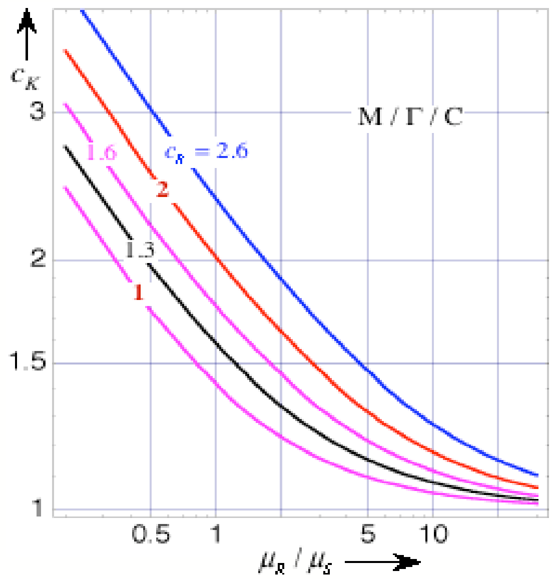

Fig. 2. Coefficient of variation for K subject to Gamma distributed residence times versus mobility ratio.

Now, we terminate the M/G case studies by the M/D/C model. This case is of special interest because it models the number of joint (convolved) integer-valued intervals, e.g. T_0 which the session time S may contain. Here

$$f_R(x) = \delta(x - T_0),\ m_{1R} = T_0,\ c_R \equiv 0, \text{ and}$$

$$p = f_R^*(s) = e^{-sT_0}.$$

Then we get for the M/D/C scenario with $\kappa = (\mu_R / \mu_S) = 1 / T_0 \mu_S$ the coefficient of handoff variations

$$c_K = \left[\frac{1}{\kappa} \frac{1 + e^{-1/\kappa}}{1 - e^{-1/\kappa}} - 1 \right]^{\frac{1}{2}} = \left[\frac{1}{\kappa} Coth\left(\frac{1}{2\kappa} \right) - 1 \right]^{\frac{1}{2}}. \tag{11}$$

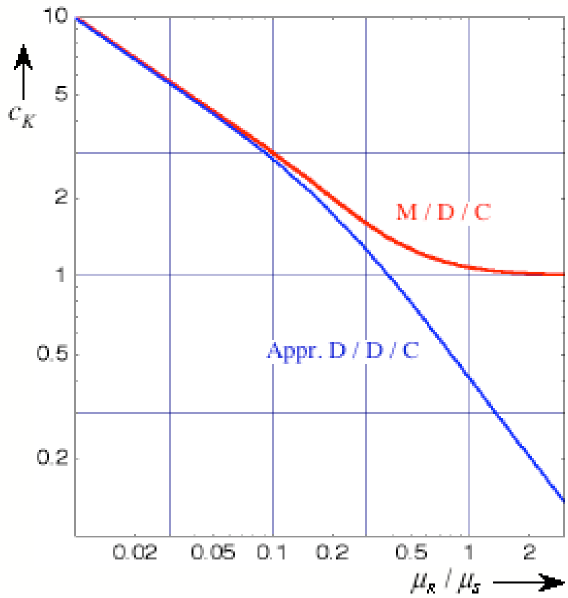

Fig. 3. Coefficient of variation for K subject to deterministic residence time distributions with differnt distributed seession times versus mobility ratio.

Now, **Fig. 3** depicts c_K versus κ for the underlying system (red). We see again that for sufficient large κ the variation of handoffs $c_K \to c_S = 1$. A surprising characteristic occurs by the further plots approximately given by $\tilde{c}_K = \sqrt{c_k^2 + (c_S^2 - 1)}$ (blue) where $c_S = 0$ defines the approximation for a D/D/C system. The agreement of both plots for small κ indicates the dominant influence of $\tilde{R}$ for low mobility ratios completed by $\tilde{c}_K \to c_S$ for large κ.

4. Generic Generalizations and the Handoff Capacity of a Cluster or a Session

Here we proceed again stepwise to generic G/G/C systems. Those require the solution of

$$P\{K > k\} = \int_0^\infty \bar{F}_S(x) g_R(x,k)dx = \int_0^\infty \bar{F}_S(x) L^{-1}\{g_R^*(s,k)\}dx\,. \tag{12}$$

Several methods of different complexities exist to solve this integral. We maintain our hybrid original-transform domain solution expressed by the right-hand side of (12). This problem decomposition gives the following advantages. First, $g_R(x,k)$ exists for many types of delayed renewal processes and saves the pure product form of the integrant. Second, we always obtain physically transparent performance interpretations. Third and in contrast to a complete transform domain solution, complicated Laplace transforms of the interrelating processes are bridged with a significant smaller amount of derivations.

For an exemplary specification of the residence time we maintain the Gamma distribution which may be matched to real field data by two moments. So we have the pair of the resulting residence time distribution density conditioned on the handoff state k

$$g_R^*(s,k) = \frac{1 - f_R^*(s)}{m_{1R}s}[f_R^*(s)]^k = \mu_R \frac{1 - (\frac{\mu_R\gamma}{s+\mu_R\gamma})^\gamma}{s}\left(\frac{\mu_R\gamma}{s+\mu_R\gamma}\right)^{\gamma k} \text{ and} \tag{13a}$$

$$g_R(x,k) = \mu_R\left\{\frac{\Gamma[(1+k)\gamma, x\mu_R\gamma]}{\Gamma[(1+k)\gamma]} - \frac{\Gamma[k\gamma, x\mu_R\gamma]}{\Gamma[k\gamma]}\right\} \text{ respectively.} \tag{13b}$$

Keying session time generalizations include hyper-exponential distributions with $c_S^2 \geq 1$, deterministic distributions with $c_S^2 \equiv 0$ and Gamma distributions $0 < c_S^2 \leq 1$ again. For hyper-exponential distributions (H) we assume an exemplary two component CDF denoted by H_2

$$\bar{F}_S(x) = \beta e^{-\mu_{S1}x} + (1-\beta)e^{-\mu_{S2}x}\,. \tag{14}$$

Unfortunately, the matching by two moments of field data requires three equations for the three unknown $(\beta, \mu_{S1}, \mu_{S2})$, [10]. Thus and for brevity we prefer here the balanced mean version defined by $\mu_{S1} = 2\beta\mu_S$, $\mu_{S2} = 2(1-\beta)\mu_S$. Now we have

$$m_{1S} = \frac{\beta}{\mu_{S1}} + \frac{1-\beta}{\mu_{S2}} = \frac{1}{\mu_S}\,,\; m_{2S} = 2\left[\frac{\beta}{\mu_{S1}^2} + \frac{(1-\beta)}{\mu_{S2}^2}\right],\; c_S^2 = \frac{m_{2S}}{m_{1S}^2} - 1\,. \tag{15a}$$

Solving the last equation subject to the balance conditions yields

$$\beta_1 = \frac{1}{2}\left(1 + \sqrt{\frac{C_S^2 - 1}{C_S^2 + 1}}\right),\; \beta_2 = 1 - \beta_1,\; c_S^2 \geq 1\,. \tag{15b}$$

These two roots form a horizontal parabola with two branches originating from ½ at $c_S = 0$ and asymptotically approaching one or zero for large c_S. One can select one branch of β only because the other branch is already considered by its complement $(1-\beta)$.

Now, we use our simple basics of chapter 3 for an enhanced statistical characterization of the assumed $\mathrm{H_2/\Gamma/C}$-system. Thus, the generalization of (14b) gives

$$P\{K>k\} = \frac{\kappa}{2}\left[(1-p_1)p_1^k + (1-p_2)p_2^k\right],\, p_i = f_R^*(\mu_{Si}) = \left(\frac{\mu_R\gamma}{\mu_{Si}+\mu_R\gamma}\right)^{\gamma}. \qquad (16)$$

$$m_{1K} \equiv \kappa\,,\; m_{2K} = \sum_{k=0}^{\infty}(2k+1)P\{K>k\} = \frac{\kappa}{2}\left\{\frac{1+p_1}{1-p_1}+\frac{1+p_2}{1-p_2}\right\}.$$

Here we have to substitute

$$p_1 = \left(\frac{\kappa\gamma/2\beta}{1+\kappa\gamma/2\beta}\right)^{\gamma},\; p_2 = \left(\frac{\kappa\gamma/2(1-\beta)}{1+\kappa\gamma/2(1-\beta)}\right)^{\gamma},\; \beta = \frac{1}{2}\left(1+\sqrt{\frac{C_S^2-1}{C_S^2+1}}\right),\; \gamma = c_R^{-2},\; \kappa = \frac{\mu_R}{\mu_S},$$

and obtain the squared coefficient of variation for the number of residence intervals within a non interrupted session the exact symbolic generalization of (7c)

$$c_\kappa^2(\kappa, c_S, c_R) = \frac{m_{2K}}{m_{1K}^2} - 1 = \frac{1}{2\kappa}\left(\frac{1+p_1}{1-p_1}+\frac{1+p_2}{1-p_2}\right) - 1\,.\; (\mathrm{H_2/\Gamma/C}) \qquad (17)$$

Its full form down to each of the parameter triple (κ, c_S, c_R) is enhanced irrational and provided in the Appendix A1. First of all (17) yields simple symbolic forms in the following special cases

$$c_\kappa^2(\kappa, c_S, c_R) = \begin{cases} 1+\kappa^{-1}, & (\mathrm{M/M/C}), \\ c_S^2+\kappa^{-1}, & (\mathrm{H_2/M/C}), \\ \frac{1}{\kappa}\mathrm{Coth}(\frac{1}{2\kappa})-1, & (\mathrm{M/D/C}). \end{cases} \qquad (18)$$

Note, that all Markovian session cases agree with (10) and (11) respectively. Now by (17) and the first case of (18) we may propose the estimate for the entire range of c_S

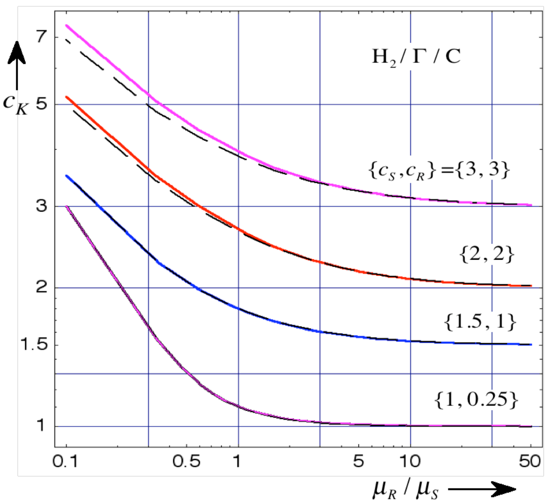

Fig. 4. Coefficient of variation for K for hyper-exponential distributed session times and Gamma distributed residence times versus mobility ratio; dashing patterns: Approximation.

$$\hat{c}_\kappa^2(\kappa, c_S, c_R) = c_\kappa^2(\mathrm{M/\Gamma/C}) + (c_S^2-1) = \frac{1}{\kappa}\frac{1+H^\gamma}{1-H^\gamma} - 1 + (c_S^2-1),\; H = \frac{\kappa\cdot\gamma}{1+\kappa\cdot\gamma},\; \gamma = c_R^{-2}. \qquad (19)$$

Now, **Fig. 4** depicts selected case studies for the underlying $\mathrm{H_2/\Gamma/C}$ system including those of the proposed approximation. In essence we have a small mobility term $\sim c_R^2\cdot\kappa^{-1}$ and a high mobility limit $\sim c_S^2$ and obtain as a crude rule of thumb for superficial discussions by

$\tilde{c}_{\kappa}^2(\kappa, c_S, c_R) \sim (c_R^2/\kappa) + c_s^2$. This behavior corresponds to the integral in the original domain form. There, the conditional mean of $g_R(x,k)$ is proportional to $m_{1R}k$ and the moments of the session time survive exclusively for large mobility ratios $\kappa = (\mu_R/\mu_S) = m_{1K}$. Moreover, comparable limits may be expected for all types of session time distributions with finite moments. But unfortunately, the process-interference for small κ remains non-intuitive. So, we must attack cases of $c_S^2 \to 0$ in order to extract the pure residence time behavior.

Therefore we proceed to a $\mathrm{D}/\Gamma/\mathrm{C}$ system. First, we have

$$\bar{F}_S(x) = 1 - U(x - T_0) = \bar{U}(x - T_0),\ m_{1S} = 1/T_0,\ c_S \equiv 0 .$$

$$P\{K > k\} = \int_0^\infty \bar{F}_S(x) g_R(x,k)dx = \int_0^{T_0} g_R(x,k)dx$$

Fortunately this integral exists and we get with $\kappa = \mu_R T_0$, $\gamma^{-1} = c_R^2 > 0$ and abbreviated expressions *E, F* respectively

$$P\{K > k\} = 1 + \kappa E - \gamma^{-1} F,\quad (\mathrm{D}/\Gamma/\mathrm{C}) \tag{20}$$

$$E = \frac{\Gamma[(k+1)\gamma, \kappa\gamma]}{\Gamma[(k+1)\gamma]} - \frac{\Gamma[k\gamma, \kappa\gamma]}{\Gamma[k\gamma]},\ \text{and}\ F = \frac{\Gamma[(k+1)\gamma + 1, \kappa\gamma]}{\Gamma[(k+1)\gamma]} - \frac{\Gamma[1 + k\gamma, \kappa\gamma]}{\Gamma[k\gamma]}.$$

Finally, the desired statistical parameters are

$$m_{2K} = \sum_{k=0}^{k_0} (2k+1) P\{K > k\}\ \text{and}$$

$$c_K^2 = \frac{m_{2K} - m_{1K}^2}{m_{1K}^2} = \frac{m_{2K}}{\kappa^2} - 1,\ \ m_{1K} = \kappa > 0 .$$

Unfortunately, the sum must be evaluated numerically with sufficient large $k_0 < \infty$ determined by the desired accuracy. Now, the before mentioned approximation can be checked again. Indeed, for $c_S = 0$ the $\mathrm{D}/\Gamma/\mathrm{C}$ system should be approximated by

$$\hat{c}_{\kappa}^2(\kappa, 0, c_R) \cong c_{\kappa}^2(\kappa, 1, c_R) - 1 . \tag{21}$$

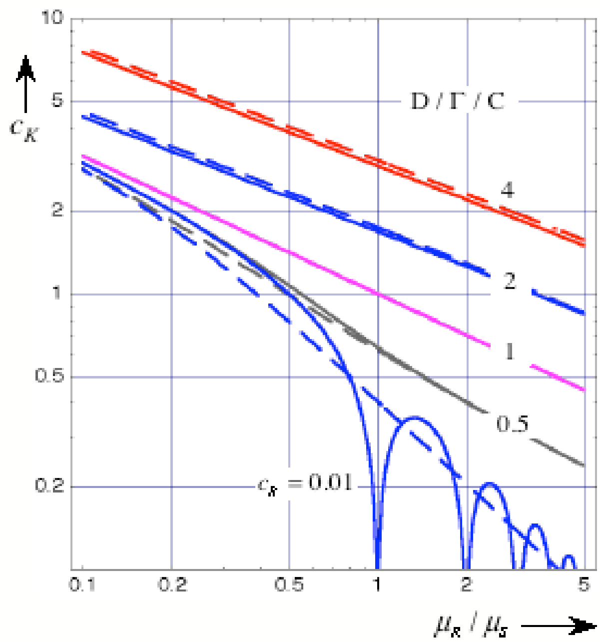

Fig. 5. Coefficient of variation for *K* subject to deterministic session time distributions with Gamma distributed residence times versus mobility ratio.

Fig. 5 compares the numerical solution with the proposed approximation (dashed lines) for different $c_R = \{0.01, 0.5, 1, 2, 4\}$. The equality sign holds for $c_R = 1$ but for small varia tions the system tends to a D/D/C one with periodic low variations of *K* if *R* divides *S.* And yet, the approximation overcomes the singularities with acceptable prediction quality.

Furthermore the CDF of a $D/\Gamma/C$ system exists in the symbolic Form of (20) and is depicted in **Fig. 6** subject to selected values of c_R. Note that we omit here and elsewhere the real staircase shapes because the resulting plots would frequently penetrate each other. Then al shapes signify lower bounds of exact values for integer *K*.

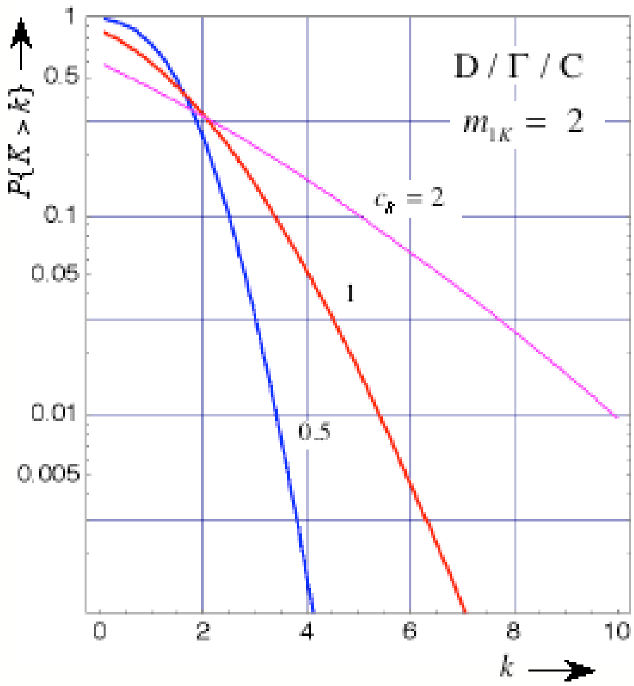

Fig. 6. CDF or survival probability of K versus k for a D/Γ/C system.

Finally we consider the generic general case of a $\Gamma/\Gamma/C$ system. Here, the advantage of the hybrid approach shortcuts the complexity enhancement of any complete transform domain approach as seen e.g. in [4 - 5]. Both alternatives terminate with numerical integrations but we remain shortly, gain physical transparency and avoid extensive branch cut considerations in the complex plane. Thus, we solve (12, 13) and define at first the session oriented Gamma CDF by

$$\bar{F}_S(x) = \int_{\alpha x}^{\infty} \frac{\alpha(\alpha y)^{\gamma-1} e^{-\alpha y}}{\Gamma(\gamma)} dy = \int_{\alpha x}^{\infty} \frac{(u)^{\gamma-1} e^{-u}}{\Gamma(\gamma)} du = \frac{\Gamma(\gamma,\ \alpha x)}{\Gamma(\gamma)}. \tag{22}$$

Since $m_{1S} = (\gamma/\alpha) = 1/\mu_S$ and $\sigma_S^2 = (\gamma/\alpha^2) = \gamma^{-1}/\mu_S^2$ we have $\alpha = \gamma \cdot \mu_S$. and $c_S^2 = 1/\gamma$. Now we distinct between the scale parameters $\{\alpha_S, \alpha_R\}$ and shape factors $\{\gamma_S, \gamma_R\}$ and get the numerical solution by the substitution of (13b) and (22) in (12) and get

$$P\{K > k\} = \bar{F}_K(k, \mu_S, \mu_R, c_S, c_R) = \int_0^{\infty} \bar{F}_S(x) g_R(x,k) dx. \tag{23}$$

Due to the throughout implied Gamma function components this integral runs excellently stable and fast provided that c_S and c_R are both not too small. A compact visualization requires data reduction. So we set $(\mu_R/\mu_S) = \kappa = 2$ and depict the number of handoffs CDF in **Fig. 7** for selected process combinations and coefficients of variation $\{c_R, c_S\}$. The Γ/M/C $(c_S = 0.01)$ and D/M/C results can't be distinguished. The further Γ/Γ/C plots imply c_S including $c_S = 1$ for easier comparisons. Furthermore we depict H_2/Γ/C and Γ/Γ/C both for $c_S = c_R = 1.8$. Indeed, both shapes differ slightly including breakpoints because the H_2/Γ/C case implies session times of different aging rates namely $(\mu_{S1}/\mu_{S2}) = [\beta/(1-\beta)] = 6.32$.

Since, the Γ/Γ/C system covers the wider range of $\{c_S, c_R\}$ combinations we prefer this as a valuable reference. Thus, the next three figures depict the new call or session handoff probability $P\{K>0\}$ and conditional handoff probability $P\{K>k \mid K>k-1\}$ respectively, cf. the Appendix A 1.

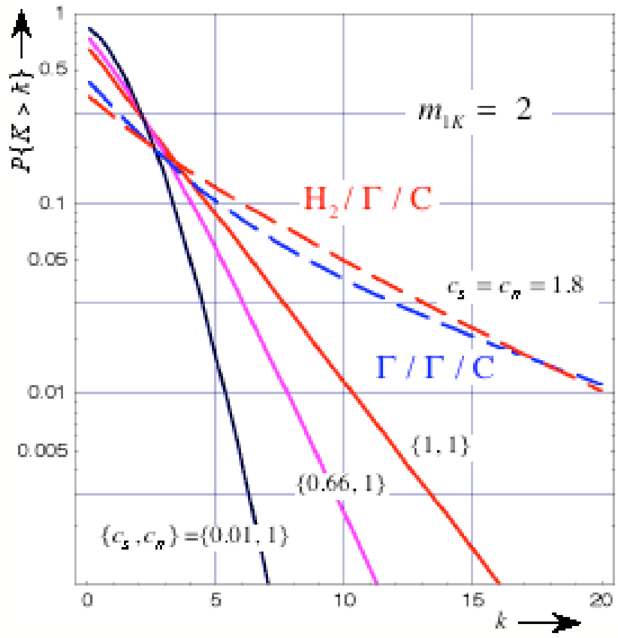

Fig. 7. CDF or survival probability of K versus k for a Γ/Γ/C in comparison to a H_2/Γ/C system .

Our more general Γ/Γ/C scenario should cover selected E_m/M/C process combinations of [6] too and simplify arbitrary extensions to a wide variety of presentations. **Fig. 8** confirms that the conditional handoff probability subject to a constant mean of handoffs m_{1K} depends on the handoff state k if the session time distribution deviates from being exponentially.

The next two figures consider fixed k subject to varying mobility measures. For easier comparisons we use the same call-to-mobility factor $(\mu_S/\mu_R) = 1/\kappa$ and parameters $\{c_S, c_R\} = \{m^{-1/2}, 1\}$ as in [6] but proceed from the underlying E_m/M/C case to the more general Γ/Γ/C scenario.

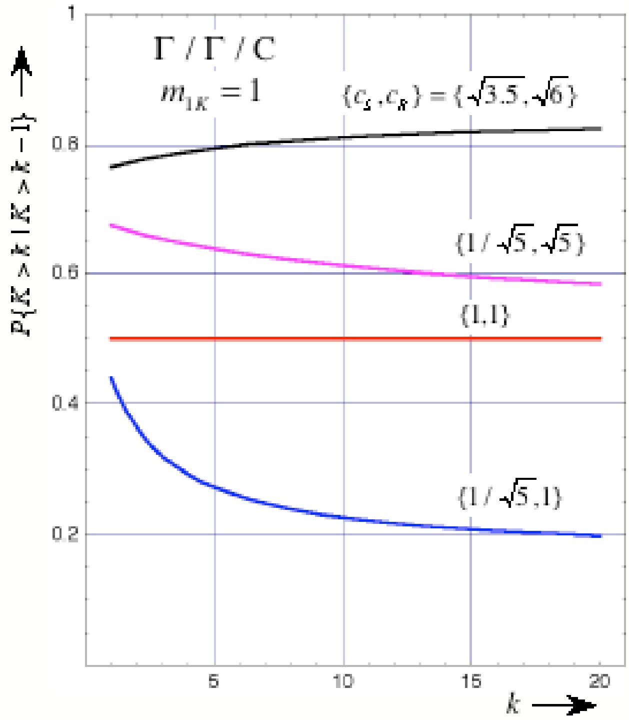

Fig. 8. The conditional handoff probability versus k .

Now, very significant differences occur if the residence time variation increases. This lowers $P\{K>0\}$ in **Fig. 9** but enhances the probability $P\{K>k \mid K>k-1\}$ in **Fig. 10**. Both figures show complete agreements with those of [6] if its parameters are used. Note, that non-integral values of the squared coefficients of variation are feasible. Besides Fig. 9, any prediction of the process behaviors prove to be much more transparently covered by the renewal oriented CDFs as shown in Fig. 7 where the conditional probabilities remain incorporated.

We tacitly assumed that the number of cells of a provisioning area is very large in comparison to relevant values of *k*. In essence the cell capacity of the cluster approches $C \to \infty$. This assumption avoids boundary effects that could make it very difficult to comprehend the dominant effects. In this case the residence or handoff capacity $C_S^{(R)}$ of a session remains a statistical parameter except. But in the more realistic case case, a limited number of cells in a provisioning area may be assumed and denoted by $C_C^{(C)}$.

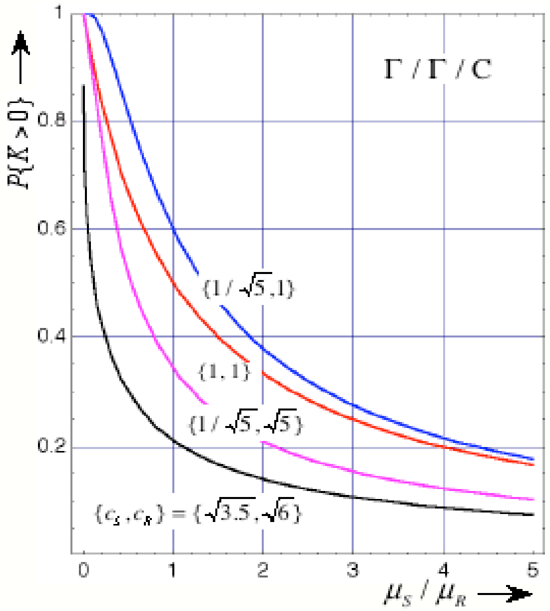

Fig. 9. The new call or session handoff probability versus the reciprocal mobility ratio.

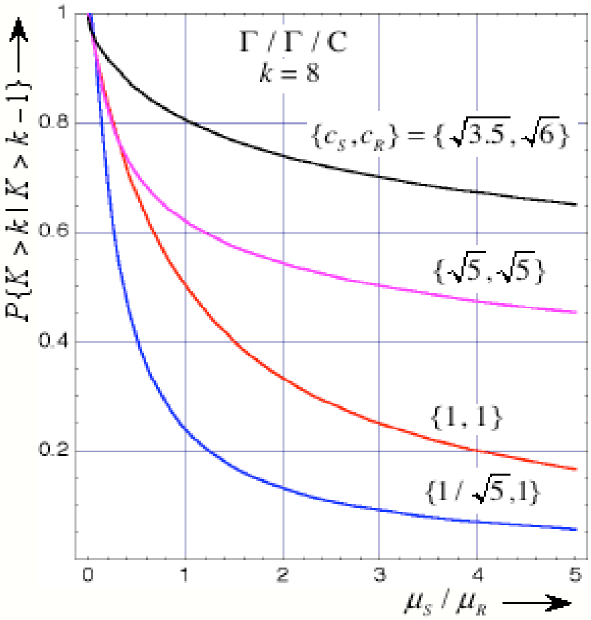

Fig. 10. The conditional call or session handoff probability versus the reciprocal mobility ratio.

Thus, we have $C = \min\{C_C^{(C)}, C_S^{(R)}\}$ and can proceed to its quantitative validation.

Here we may use either the required survival probability of a session $P_S^{(R)}$ or determine the roaming probability for the provisioning area $P_R^{(C)}$ by

$$P\{K > C_S^{(R)}\} - P_S^{(R)} = 0 \quad \text{or} \quad P_R^{(C)} = \frac{P\{K = C_C^{(C)}\}}{\sum_{j=0}^{C_C^{(C)}} P\{K = k\}} \tag{24}$$

respectively. The left hand equation requires a find-root-operation and the right hand one results by a truncation of (5a) and may be computed straightforward subject to the given number of cells $C_C^{(C)}$. The most general case is rather attackable by inequalities of different tightness. For an immediate intuitive feeling, we recall the well-known Markov inequality $P\{X > x\} \le m_{1X} / x$. Then we have respectively

$$C_S^{(R)} \ge m_{1K} / P_S^{(R)} \quad \text{or} \quad P_R^{(C)} \le \underline{m}_{1K} / C_C^{(C)}. \quad \text{(G/G/C)} \tag{25}$$

Here $\underline{m}_{1K} \cong m_{1K}$ is mean number of handoffs subject to the effective truncation. The estimates (25) uniformly cover both cases but far more tight values result by the simple operations (24).

5. Selected Applications

The following applications consider possible signaling rate estimations of the UDP-based Radius- (Remote Authentication Dial In User Service) and the TCP-based Diameter-protocols. For the former we refer to [8] and for the latter our renewal model subject to forced terminations is applied. Here we assume the possible Diameter signaling rate model similar to [9]. In any case we must count the number of intervals *I*, say interims within general distributed times and defined by the random Variable *X*. Then the mean number of interims subject to given step intervals Δx can be computed, at least numerically. But the first moment for generally distributed variables can be estimated more directly by intuitive reasoning.

For a given variable *X* of known or measured probability distribution we simply guess $\hat{m}_{1I}(\Delta x) = m_{1X} / \Delta x$. Indeed, this estimate satisfies the following upper bound relation

$$\hat{m}_{1I}(\Delta x) = \frac{1}{\Delta x}\int_0^\infty x \cdot f_X(x)dx \;\geq\; \sum_{i=0}^\infty i \int_{i\Delta x}^{(i+1)\Delta x} f_X(x)dx = m_{1I}(\Delta x) . \tag{26}$$

Now, the covered intervals define the new discrete probabilities according to

$$P_i(\Delta x) = P\{i\Delta x < X \leq (i+1)\Delta x\} = F_X[(i+1)\Delta x] - F_X(i\Delta x) = \int_{i\Delta x}^{(i+1)\Delta x} f_X(x)dx ,\; x_i \geq 0 ,\; \sum_{i=1}^\infty P_i = 1 .$$

Then, we obtain the higher moments expressed by the CDF of X

$$m_{\ell I}(\Delta x) = \sum_{i=0}^\infty i^\ell \cdot P_i(\Delta x) = \sum_{i=0}^\infty i^\ell \left\{ \bar{F}_X(i\Delta x) - \bar{F}_X[(i+1)\Delta x] \right\} \text{ or}$$

$$m_{\ell I}(\Delta x) = \sum_{i=0}^\infty i^\ell \bar{F}_X(i\Delta x) - \sum_{j=1}^\infty (j-1)^\ell \bar{F}_X[j\Delta x] \equiv \sum_{j=1}^\infty [j^\ell - (j-1)^\ell]\bar{F}_X(j\Delta x) . \tag{27a}$$

So, we obtain e.g.

$$m_{1I} = \sum_{j=1}^\infty \bar{F}_X(j\Delta x) ,\; m_{2I} = 2\sum_{j=1}^\infty \bar{F}_X(j\Delta x) - m_{1I} \text{ and } c_I^2 = (m_{2I} - m_{1I}^2) / m_{1I}^2 . \tag{27b}$$

In essence the underlying aggregation needs at least one interval Δx and their moments differ basically from those of ordinary variables considered in Chapter 3.

We select again the Gamma probability distribution as a generic reference. Its well-known CDF is

$$\bar{F}_X(x,\gamma) = \int_x^\infty \frac{\alpha^\gamma y^{\gamma-1} e^{-\alpha y}}{\Gamma(\gamma)} dy = \int_{\alpha x}^\infty \frac{t^{\gamma-1}e^{-t}}{\Gamma(\gamma)} dt = \frac{\Gamma(\gamma,\alpha x)}{\Gamma(\gamma)} . \tag{28}$$

α is called the scale parameter and $\gamma = 1/c_X^2$ the shape factor. $\Gamma(\gamma,\alpha x) = \int_{\alpha x}^\infty t^{\gamma-1}e^{-t}dt$ is the incomplete Gamma function and $\Gamma(\gamma,0) \equiv \Gamma(\gamma)$ the Euler Gamma function. Furthermore, for integer $\gamma = r$ and thus $\Gamma(\gamma) = (r-1)!$ the Erlang-*r* CDF results. It models the accumulated service of *r* stages for one initially exponential distributed job in a server chain according to

Burke's theorem. The coefficient of variation c_X signifies that $0 < \gamma < \infty$ covers a wide range of dispersions including the NED case where $\gamma \equiv 1$. So the mean number of incremental intervals Δx within the assumed distribution X becomes

$$m_{1I}(\Delta x) = \sum_{i=1}^{\infty} \bar{F}_X(i\Delta x) = \sum_{i=1}^{\infty} \int_{\alpha i \Delta x}^{\infty} \frac{t^{\gamma-1} e^{-t}}{\Gamma(\gamma)} dt \,. \tag{29}$$

Thus, we may substitute $\alpha = \mu_X \gamma$ and $\gamma = c_X^{-1/2}$. Then, plots for selected values of c_X, in essence, the mean number of intervals $m_{1I}(\Delta x)$ versus $\mu_X \Delta x$ are shown in **Fig. 11.** Indeed, our estimate $\hat{m}_{1I}(\Delta x)$ indicates the upper bound. The NED variant with $c_X = 1$ appears here as a special case but highlights the somewhat complicated summing operation in (29) by

$$m_{1X}(\Delta x) = \sum_{i=1}^{\infty} \int_{i\mu_X \Delta x}^{\infty} \frac{e^{-t}}{\Gamma(\gamma)} dt = \sum_{i=1}^{\infty} e^{-i\mu_X \Delta x} = \frac{e^{-\mu_X \Delta x}}{1 - e^{-\mu_X \Delta x}} \,.$$

The further non-algebraic explicit forms of (29) for $c_X \neq 1$ became evaluated numerically. For these cases, the sum was limited to a finite number of components for sufficient high accuracy. Note, that a high variation of X justifies the upper bound approach also.

Now, we proceed to the total mean signaling rates and emphasize that their components may change with any protocol interpretation but doesn't affect the handoff fundamentals presented throughout this contribution.

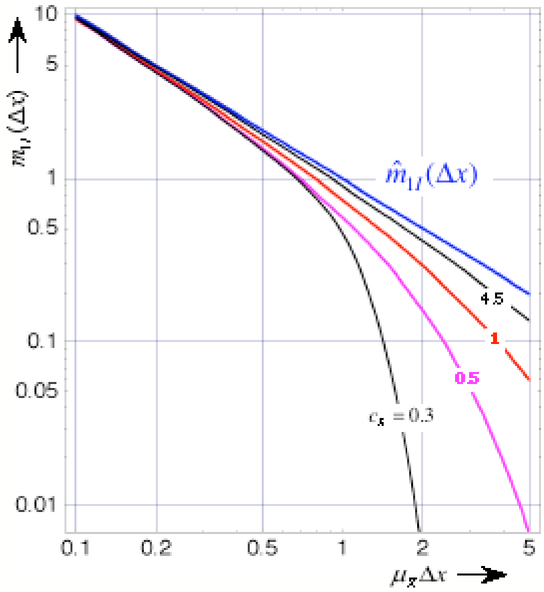

Fig. 11. Upper bound and mean values of aggregated Gamma probability density versus the normalized aggregation interval.

First, the Radius rate (*RR*) has been considered in [8] and composed by the following components: Initial authentication, start/stop respectively plus the mean number of interims. Here we avoid different acceptance factors by the consideration of service dependent blockings $P_b^{(j)}$. Then mean Radius rate generated by each Packet Data Serving Node (PDSN) i and type of service j may be estimated by

$$E\{RR\} = \sum_{i=1}^{N_{\mathrm{PDSN}}} \sum_{j=1}^{N_{\mathrm{S}}} \lambda^{(i,j)} \left\{1 + [2 + m_{1I}^{(j)}(\Delta T^{(j)})](1 - P_b^{(i,j)})\right\}. \tag{30}$$

Note that no handoff traffic occurs with Radius.

Next, we derive a Diameter rate formula which differs from early suggestions in [9]. Basically there are six components, namely initial authentication, accounting request, start/stop, handoffs (*K*) interims (*I*) and authorization lifetimes (*M*). Early publications [1], [2] show that call or session blocking remains a local cell-control oriented GoS parameter. It depends on the arrival and departure process including the appropriate cell queueing model, seen e.g. in [7]. Fortunately, the channel holding time of each cell remains locally cell-oriented also, leaving the session time of our trajectories unchanged. But as a consequence of blocking forced terminations of sessions may occur and these change the session distribution function. Consequently the expected value of handoffs changes too. So we are faced with different blocking impacts, in essence the local blocking for an arriving session P_b and its accumulated value caused by possible handoffs P_f. Furthermore, the mean of completed handoffs m_{1Kf} subject to forced terminations and finally the change of interims due to forced terminations $m_{1If}(\Delta x)$ are to be considered. For lack of space we show in the Appendix A3 that our hybrid approach may be easily extended to any accumulating handoff blocking and this becomes important because the cell sizes decrease if the bandwidth requirements continue to increase.

In comparison to (30) we avoid a huge amount of possible parameter combinations by the sum of all source-specific Poisson arrivals $\Lambda = \sum_{i=1}^{Np}\sum_{j=1}^{Ns}\lambda^{(i,j)}$ together with the average values $m_{1X} = \frac{1}{\Lambda}\sum_{i=1}^{Np}\sum_{j=1}^{Ns}\lambda^{(i,j)}m_{1X}^{(j)}$ for all first moments of the variable X under consideration, $P_b = \frac{1}{\Lambda}\sum_{i=1}^{Np}\sum_{j=1}^{Ns}\lambda^{(i,j)}P_b^{(i,j)}$ and $P_f = \frac{1}{\Lambda}\sum_{i=1}^{Np}\sum_{j=1}^{Ns}\lambda^{(i,j)}P_f^{(i,j)}$. So and supported by suggestions from [9] we may compose the mean Diameter rate estimate by

$$E\{DR\} = \Lambda\left\{\{1+(1-P_b)[1+(1-P_f)]\}(1+m_{1Kf}) + [m_{1I}^{(T)}(\Delta T) + m_{1I}^{(T)}(\Delta M)]\right\}. \tag{31}$$

Here T denotes the effective session time the distribution of which differs from that of S in general if $p_f \neq 0$. Observe that this relation generates the mean Radius rate according to (30) for vanishing P_f, m_{1Kf} and $m_{1I}^{(T)}(\Delta M)$. A rigorous derivation of these functions is provided in [12] but verified in the Appendices A2 – A3. Here we refer to the M/G/C case and have

$$P_f = \frac{P\{K>0\}p_f}{1-p(1-p_f)},\ m_{1Kf} = \frac{m_{1K}(1-p)}{1-p(1-p_f)} \text{ and } m_{1I}^{(T)}(\Delta x) \cong \frac{e^{-\mu_S(1+p_f m_{1K})\Delta x}}{1-e^{-\mu_S(1+p_f m_{1K})\Delta x}}. \tag{32}$$

Comments: All equations reduce to intuitive base cases for $p_f = 0$. The first equation states that the accumulated probability concerns at least one handoff. Next, forced terminations decrease the mean number of handoffs because $m_{1Kf} \leq m_{1K}$ and this corresponds to an increased session completion rate as seen by the last equation.

Finally, we may consider uneconomic over-provisioning of the cell switching capabilities so that all blocking components vanish and $m_{1Kf} \equiv m_{1K}$, and the upper bounds of $m_{1I}(\Delta x)$ respectively. Then we obtain the generic upper bound of the mean DR by

$$\hat{E}\{DR\} = \Lambda\left\{3\left(1+m_{1K}\right)+\left[\hat{m}_{1I}(\Delta T)+\hat{m}_{1I}(\Delta M)\right]\right\}. \text{ (G/G/C)} \quad (33)$$

Now, **Fig. 12** depicts (30) and (31) versus m_{1R} / m_{1S} for two mean values of total arrival rate Λ. First of all the dashed lines are the G/G/C upper bounds resulting from (33). Next, we observe within the plots of (31) the changing sensitivity with respect to forced terminations and mobility. In essence, for high mobility values significant rate reductions compete with increasing session droppings. Further experiences with forced terminating M/G/C systems show that both values m_{1Kf} and $m_{1I}^{(T)}(\Delta x)$ may be approximated by M/M/C formulas because the effect of residence time variations remains very small. Moreover, it is derived in [12] that the probability density distribution of the effective session time T in this case preserves its exponential form but changes its moments. For example

$$f_T(t) \cong \mu_S(1+p_f\kappa)e^{-\mu_S(1+p_f\kappa)t}\,,\ \kappa = \frac{\mu_R}{\mu_S} \equiv m_{1K}\,. \quad (34)$$

Note that $f_T(t) \equiv f_S(t)$ for $p_f = 0$. Furthermore and with respect to G/G/C scenarios under progress the effective session distribution T changes mainly due to the original session time S and much more less due to that of R as seen by the variation dependencies in Chapter 4.

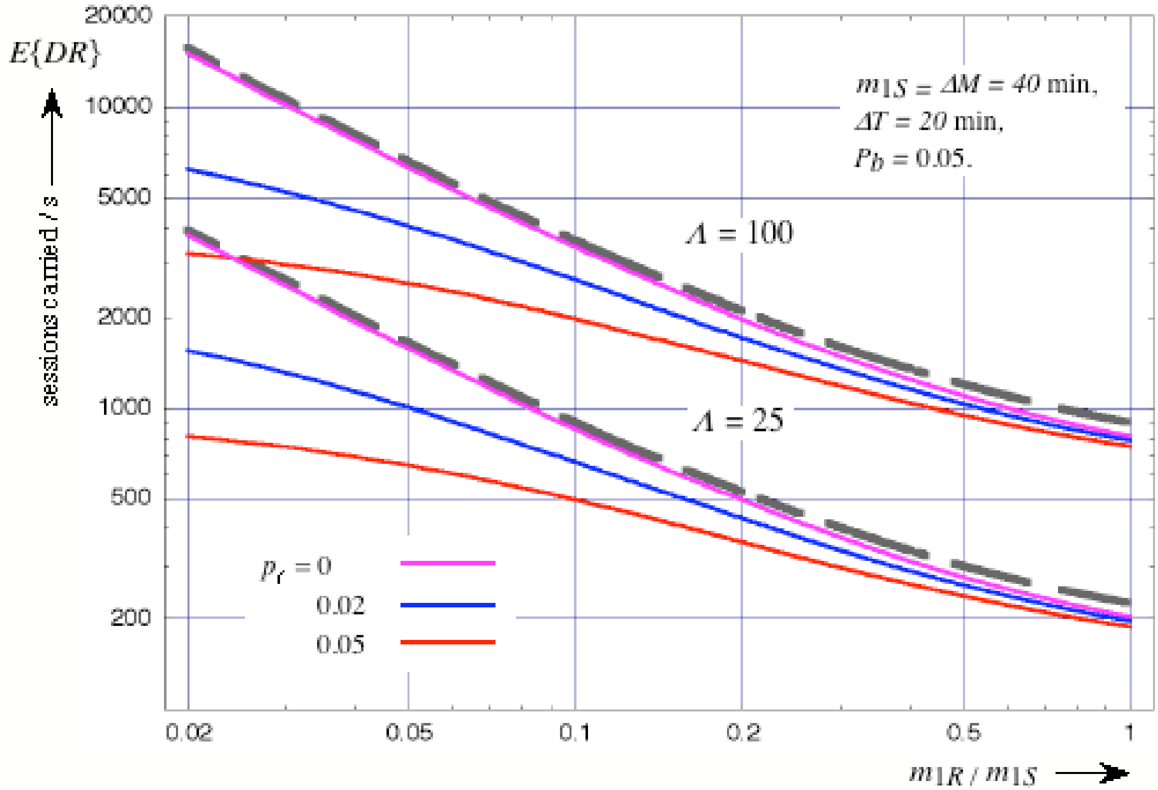

Fig. 12. Upper bounds (dashed shapes) of mean Diameter rates for G/G/C systems in comparision to M/M/C ≈ M/Γ/C scenarios with arbitrary γ $(0.5 \le \gamma \le 2)$ and selected forced termination probabilities p_f versus m_{1R} / m_{1S}.

Conclusions

The handovers of each session in cellular networks are modeled by a delayed renewal process with possible defections which provides the main statistical parameters of the counted handoffs very compactly. The well-known pure transform-domain solutions loose much of any desired physical transparency in view of the simplification problem and the page consuming relations. In essence early contributions of Hilbert, Matijasevich and Roberts showed that the simplification problems can‘t be solved by algorithms. We avoid this by a hybrid original-transform-domain approach. This directly proceeds to operational general process interrelations session/residence subject to the provisioning area (cluster) capacity. It is shown that the mean number of handoffs are independent of the process inter-occurrence time distributions if handoff blockings can be ignored. Under such over-provisioning conditions relevant second order statistics ranging from $\mathrm{H}_2/\Gamma/C$ to more general $\Gamma/\Gamma/C$ scenarios are obtained. Both types of densities may be easily fitted on real applications by two equations for each dedicated process component only. For instance, hyper-exponential session time distributions, modeling IP mixtures of interactive short message and large file transfers are covered too. A new equation which interrelates the handoff's coefficient of variation with those of session and residence times shows that the session variation dominates for high mobility ratios and the residence variation determines the handoff variation for low mobility ratios. Moreover the complementary handoff distribution functions and their single step conditional probabilities are derived. These functions allow to monitor a session in progress and plan ahead the next handoff of the session. The provided fundamentals allow their extension to new and handoff session blockings. Then, the expected value of handoff and session time changes. Including this, the mean signaling rate of AAA procedures may be calculated where the mean number of handoffs and mean effective session time determine the expectable arrival rates of the AAA server.

The advantages obtained here are the simplicity of the premises, the permissibility of general distributed interrelating processes and a wide range of possible refinements. Until the time of writing the arrived hybrid approach covered the incomplete and complete session probabilities and the resulting effective session time subject to forced terminations but these performance samples were focused on M/G/C cases only. Starting from the provided relations, methodical similar evaluations with respect to generic G/G/C systems can be executed. Finally, more precise signaling protocol models and accounting schemes may be considered without any change of the basic fundamentals presented here.

Acknowledgement

Thanks are devoted to Yuguang Fang for his hints with respect to the continuing importance of forced terminations and his paper [7]. The QoS order of magnitude for the blocking of new and handoff sessions mentioned in that paper agrees with those assumed in the contribution deployed here.

Appendix

A 1. Selected basic Relations

The exact symbolic c_K^2-form of $\mathrm{H_2/\Gamma/C}$ down to each of the parameter triple (κ, c_S, c_R) is enhanced irrational and reads

$$c_\kappa^2(\kappa, c_S, c_R) = \frac{1}{\kappa}(A^{-1} + B^{-1} - 1) - 1 \tag{A1a}$$

$$A = 1 - \left\{\kappa\left[-c_R^2\left(-1+\sqrt{1-\frac{2}{1+c_S^2}}\right)+\kappa\right]^{-1}\right\}^{1/c_R^2}, \quad B = 1 - \left\{\kappa\left[c_R^2\left(1+\sqrt{1-\frac{2}{1+c_S^2}}\right)+\kappa\right]^{-1}\right\}^{1/c_R^2} \tag{A1b}$$

$$c_\kappa^2(\kappa, c_S, 1) = c_S^2 + 1/\kappa, \quad \underset{c_R \to 0}{Limit}\, c_\kappa^2(\kappa, 1, c_R) = \mathrm{Coth}(\tfrac{1}{2\kappa}) - 1, \; c_S^2 \geq 1,$$

Unfortunately, the $\mathrm{H_2/\Gamma/C}$-formula covers the parameter range $c_S^2 \geq 1$ only. Furthermore the solution for the $\mathrm{D/\Gamma/C}$-System has no closed symbolic form. But the further special cases of $\mathrm{M/\Gamma/C}$

$$c_\kappa^2(\kappa, 1, c_R) = \frac{1}{\kappa}\frac{1+H^\gamma}{1-H^\gamma} - 1, \; H = \frac{\kappa \cdot \gamma}{1+\kappa \cdot \gamma}, \text{ with } \gamma = c_R^{-2} \tag{A2}$$

favors its heuristic extension to the entire area $0 \leq c_S^2 < \infty$ by

$$\hat{c}_\kappa^2(\kappa, c_S, c_R) \cong c_\kappa^2(\kappa, 1, c_R) + (c_S^2 - 1). \tag{A3}$$

The equality sign of this approximation then holds for $c_S^2 = 1$. Otherwise the underlying coefficient variation is corrected by the remaining complement. Surprisingly enough that moderate deviations occur only for extreme session variations as discussed in Chapter 4.

The conditional handoff probability $P\{K > k \mid K > k-1\}$ may be derived from the following definitions. First, assume $B = \{K > k\}$ and $A = \{K > k-1\}$. Thus, $A \supset B$ and by the definition $P\{B \mid A\} = P\{A \cap B\} / P\{A\}$ we get

$$P\{B \mid A\} = \frac{P\{B\}}{P\{A\}} = \frac{P\{K > k\}}{P\{K > k-1\}} \equiv P\{K > k \mid K > k-1\}. \tag{A4}$$

Thus, for each M/G/C system, according to (6), $P\{K > k \mid K > k-1\} \equiv p$ remains independent of k, see Fig. 8 too.

A 2. The Change of the Session Time due to forced Terminations

First, the modified session time of accepted calls may be decomposed into two components. As in [4] we distinct between incomplete and complete parcels. The former refer to handoff sessions subject to forced terminations due to handoff failures and the latter to all successfully terminating sessions. Within the following model the normalization $P_i + P_c + P_b = 1$ occurs implicitly.

With respect to an incomplete session we can introduce the combinations of session droppings into (6a) valid for M/G/C scenarios. Then we get for an accepted session subject to handoff blockings the accumulated forced termination probability

$$P_f = \sum_{k=1}^{\infty}\sum_{j=1}^{k}\binom{k}{j} m_{1K}(1-p)(p^{k-1}-p^k)p_f^j(1-p_f)^{k-j} \text{ or}$$

$$P_f = P\{K>0\}\sum_{k=1}^{\infty}[p(1-p_f)]^{k-1}p_f = \frac{P\{K>0\}p_f}{1-p(1-p_f)} = \frac{P_i}{1-P_b}. \quad \text{(M/G/C)} \quad \text{(A5a)}$$

Obviously, the session succeeds in each of the first $(k-1)$ handoff attempts which it requires and fails in the kth. For clearness we can determine the accumulated survival probability

$$P_g = P\{K=0\} + \sum_{k=1}^{\infty}\sum_{j=0}^{0}\binom{k}{j} m_{1K}(1-p)(p^{k-1}-p^k)p_f^j(1-p_f)^{k-j}$$

or

$$P_g = P\{K=0\} + P\{K>0\}\sum_{k=1}^{\infty}(p^{k-1}-p^k)(1-p_f)^k \equiv 1-P_f = \frac{P_c}{1-P_b} \quad \text{(M/G/C)} \quad \text{(A5b)}$$

which confirms the probability of all remaining sessions not forced to terminate. In essence $P_i + P_c = 1 - P_b$.

Now, the total probability density $f_T(t)$ of the effective session time T is

$$f_T(t) = \frac{P_i}{1-P_b}f_{Ti}(t) + \frac{P_c}{1-P_b}f_{Tc}(t) = \tilde{f}_{Ti}(t) + \tilde{f}_{Tc}(t), \qquad \frac{P_i}{1-P_b} + \frac{P_c}{1-P_b} = 1. \quad \text{(A6)}$$

For the incomplete sessions we may directly proceed to the weighted residence time density

$$\tilde{f}_{Ti}(t) = \sum_{k=1}^{\infty} g_R(t,k-1)(1-p_f)^{k-1}p_f\bar{F}_S(t) = \sum_{j=0}^{\infty} L^{-1}\left\{\frac{\mu_R}{s}(1-p)[p(1-p_f)]^j\right\}p_f\bar{F}_S(t)$$

which has the compact sum form

$$\tilde{f}_{Ti}(t) = L^{-1}\left\{\frac{\mu_R}{s}\frac{(1-p)p_f}{1-(1-p_f)p}\right\}\bar{F}_S(t). \quad \text{(A7a)}$$

For brevity we assume $p = f_R^*(s) = [\gamma_R\mu_R / (s+\gamma_R\mu_R)]^{\gamma_R}$, in essence Gamma distributed residence times and obtain the Markovian base case

$$\tilde{f}_{Ti}(t) = L^{-1}\left\{\frac{\mu_R p_f}{s+\mu_R p_f}\right\}\overline{F}_S(t) \equiv \mu_R p_f e^{-\mu_R p_f t}\overline{F}_S(t) \ , \ \gamma_R \equiv 1\,. \quad \text{(G/M/C)} \qquad \text{(A7b)}$$

This result forms an important reference for the generic general case because detailed case studies validate

$$\tilde{f}_{Ti}(t) = L^{-1}\left\{\frac{\mu_R p_f}{s+\mu_R p_f}\left[1+O\left(p_f\frac{\ln\sqrt{\gamma}}{\sqrt{\gamma}}\right)\right]\right\}\overline{F}_S(t) \cong \mu_R p_f e^{-\mu_R p_f t}\overline{F}_S(t) \quad \text{(G/}\Gamma\text{/C)} \quad \text{(A8)}$$

Roughly spoken, the second error term in the bracket appears to be very small and its residual change versus s can be neglected. This is confirmed by **Fig. A 1** for the complete Argument $L^{-1}\{...\}$ of (A8) in comparison to the reference case $\gamma = 1$ versus s subject to different γ and $p_f = 0.05$. Smaller values of p_f reduce the error term in (A8). The depicted shapes show that the exponential form remains valid aside a small offset caused by the error term given above. We will see that the approximate inversion of equations like (A6) forms a indispensable condition for further far more complicated case studies.

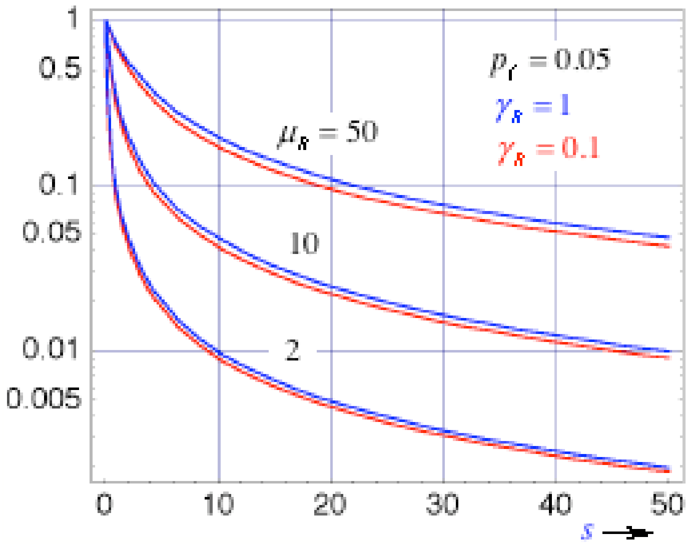

Fig. A 1. The effect of Gamma distributed residence times in comparison to exponentially distributed ones (blue) subject to forced terminations in the s-domain.

Now we turn or attention to the complete sessions with forced terminations. In this case the session remains successfully either if no handoff occurs or if the accumulated residence times survive. Then we get

$$\tilde{f}_{Tc}(t) = f_S(t)\left[\overline{F}_{\tilde{R}}(t) + \sum_{k=1}^{\infty}\int_0^t g_R(x,k-1)(1-p_f)^{k-1}(1-p_f)^k\,\overline{F}_R(t-x)dx\right]. \qquad \text{(A9)}$$

Here $\overline{F}_{\tilde{R}}(t)$ expresses the proportion of sessions for $\tilde{R} > t$ (no handoffs) so that the original session time density remains unchanged. The second sum-term includes the survival probability $(1-p_f)^k$ and the proportion of residence-excess from x until $S = t$ expressed by $\overline{F}_R(t-x)$. Similar to the derivation of (A8) we express the sum-term in the transform domain and get

$$\sum_{j=0}^{\infty} L^{-1}\left\{\frac{\mu_R}{s}(1-p)(1-p_f)[p(1-p_f)]^j\right\} = L^{-1}\left\{\frac{(1-p_f)\mu_R}{s+\mu_R p_f}\left[1+O\left(\frac{p_f}{\sqrt{\gamma}}\ln\sqrt{\gamma}\right)\right]\right\}.$$

Thus, we obtain the eminent compact near accurate approximation of (A9) by

$$\tilde{f}_{Tc}(t) \cong f_S(t)\left[\bar{F}_{\tilde{R}}(t) + \mu_R(1-p_f)\int_0^t e^{-\mu_R p_f x}\bar{F}_R(t-x)dx\right]. \quad \text{(G/Γ/C)} \qquad \text{(A10)}$$

Here the second integral-term can be symbolically computed for the generic Gamma CDF. For sake of space we maintain (A10), explain its performance in the G/M/C case but delay the exact enumeration of the Γ/Γ/C case to [12]. The former case then gives

$$\tilde{f}_{Tc}(t) \cong f_S(t)\left[e^{-\mu_R t} + e^{-p_f\mu_R t} - e^{-\mu_R t}\right] = f_S(t)e^{-p_f\mu_R t}. \qquad \text{(A11)}$$

Hence, the total probability density of the effective session time becomes

$$\tilde{f}_T(t) = \tilde{f}_{Ti}(t) + \tilde{f}_{Tc}(t) \simeq [\mu_R p_f \bar{F}_S(t) + f_S(t)]e^{-p_f\mu_R t}. \quad \text{(G/M/C)} \qquad \text{(A12a)}$$

We see immediately that $\tilde{f}_T(t) \equiv f_S(t)$ for $p_f = 0$.

For example we get

$$\tilde{f}_T(t) \equiv [e^{-\mu_s t}p_f\mu_R + \mu_s e^{-\mu_s t}]e^{-p_f\mu_R t} = [\mu_s + p_f\mu_R]e^{-(\mu_S + p_f\mu_R)t}. \quad \text{(M/M/C)} \qquad \text{(A12b)}$$

Of course, this relation holds strictly and reflects a density function. Their characteristic values are

$$m_{1T} = \frac{1}{\mu_T} = \int_0^\infty t\cdot\tilde{f}_T(t)dt = \frac{1}{\mu_S(1+p_f m_{1K})} \le \frac{1}{\mu_S} \text{ and } c_T^2 = 1. \qquad \text{(A12c)}$$

Thus, the mean effective session time decreases if the product $p_f m_{1K} = p_f\mu_R / \mu_S$ increases.

Let us consider finally higher order moments of the effective session time of G/G/C systems. There are three alternatives, namely enumeration, an original domain symbolic moment expansion and a transform domain approach. All require the definition of the underlying processes. Here we can show that the G/G/C original domain moment expansion results in an eminently transparent moment interrelation. Thus, using (A12a) we obtain the moments by

$$m_{\ell T} = \int_0^\infty t^\ell\cdot\tilde{f}_T(t)dt \cong \int_0^\infty t^\ell\cdot[\mu_R p_f\bar{F}_S(t) + f_S(t)]e^{-p_f\mu_R t}dt. \qquad \text{(A13a)}$$

Now, the expansion of the exponential function gives a keying moment generating difference equation which interrelates the higher order moments of *T* and *S*, e.g.

$$m_{\ell T} \cong \sum_{k=0}^\infty \frac{(-\alpha)^k}{k!}\left[\alpha\int_0^\infty t^{\ell+k}\bar{F}_S(t)dt + \int_0^\infty t^{\ell+k}f_S(t)dt\right] = \sum_{k=0}^\infty \frac{(-\alpha)^k}{k!}\left[\alpha\cdot m_{(\ell+k+1)S} + m_{(\ell+k)S}\right] \qquad \text{(A14b)}$$

with $\alpha = \mu_R p_f = m_{1K}p_f\mu_S$. This series converges rapidly and within $\alpha < 1$ we get e.g.

$$m_{1T} = \alpha\frac{m_{2S}}{2} + m_{1S} - \alpha[\alpha\frac{m_{3S}}{3} + m_{2S}] = m_{1S}\left(1 - p_f m_{1K}\frac{1+c_S^2}{2}\right) + O\left[\alpha^2\frac{m_{S3}}{3}\right]. \quad \text{(A14c)}$$

This is important because the squared coefficient of variation of S generalizes (A12c). Obviously (14c) declines linear versus m_{1K} but the higher moment terms cause a rounded decrease. But suggested by a comparison with (12c) and numerical studies we may avoid the more extensive higher moment terms by the useful but crude approximation

$$m_{1T} = \frac{1}{\mu_T} \simeq \frac{1}{\mu_S\left(1 + p_f m_{1K}\frac{1+c_S^2}{2}\right)} \le \frac{1}{\mu_S} \quad \simeq \text{(G/G/C)}. \qquad \text{(A14d)}$$

Besides trivial reasons for the different equality signs we see that the squared coefficient of session time variation significantly modifies the mobility dependence of the mean effective session time again assuming that residence time variation may be neglected.

A3. The mean number of Handoffs and Interims with forced Terminations

In contrast to the mean effective session time our basic moment relations (4a) and (7a) subject to forced terminations can be applied directly. G/G/C cases are to be considered in [12] but for principal understanding we focus on the M/G/C case here. For an intuitive confidence we first refer to (14b) and introduce Walds's identity. Then we proceed to the more general M/G/C case.

Let us denote the number of handoffs subject to forced terminations by the random variable K_f. Now, Wald's equation reads

$$m_{1T} = E\left\{\sum\nolimits_{j=1}^{K_f} R_j\right\} \equiv E\left\{K_f\right\}E\{R\} = m_{1K_f} m_{1R} \qquad \text{(A15)}$$

Substitution of (12c) then yields

$$m_{1K_f} = \frac{m_{1T}}{m_{1R}} = \frac{m_{1S}}{m_{1R}}\frac{1}{1+p_f m_{1K}} = \frac{m_{1K}}{1+p_f m_{1K}} \le m_{1K} \quad \text{(M/M/C)} \qquad \text{(A16)}$$

This forms a special case of more general M/G/C scenarios expressed by the following theorem.

Theorem III: The mean total number of Handoffs subject to forced terminations is given by

$$m_{1Kf} = \frac{m_{1K}(1-p)}{1-p(1-p_f)} \le m_{1K}, \; p = f_R^*(\mu_S). \quad \text{(M/G/C)} \qquad \text{(A17)}$$

Proof: The total first moment of handoffs m_{1K_f} implies two components namely those of incomplete and complete sessions m_{1Ki} and m_{1Kc} respectively. Now, appropriate completion of (7b) with $P\{K > k\} = m_{1K}(1-p)p^k$ gives

$$m_{1Ki} = m_{1K}(1-p)\sum_{k=0}^{\infty}[p(1-p_f)]^k p_f = \frac{m_{1K}(1-p)p_f}{1-p(1-p_f)}, \tag{A18a}$$

$$m_{1Kc} = m_{1K}(1-p)\sum_{k=0}^{\infty}[p(1-p_f)]^k (1-p_f) = m_{1K}\frac{(1-p)(1-p_f)}{1-p(1-p_f)}. \tag{A18b}$$

So, its sum yields (A17) and terminates the proof.

Comments: Again the dependence on the variation of residence times remains very small. For exponentially distributed residence times with $p = f_R^*(\mu_S) = [m_{1K}/(1+m_{1K})]$ we obtain (A16) again. An extension to G/G/C scenarios may be obtained similar to the suggestions given in the Subsection A2. Observe that (A17) and (A18) may be expressed also by

$$m_{1K_f} = \frac{1}{m_{1R}}\left[m_{Ti}p_f + m_{Tc}(1-p_f)\right] = \frac{1}{m_{1R}}\left[\frac{\tilde{m}_{1Ti}}{P_i}p_f + \frac{\tilde{m}_{1Tc}}{P_c}(1-p_f)\right] \tag{A19}$$

where the components $\tilde{m}_{1Tx}$ are the first weighted moments obtainable from (A10) and (A11).

Finally, our mean value experiences with forced terminations have to be considered within the interim operator (29) too. First, we may rewrite (9) and (28) by the substitution $m_{1X} \equiv m_{1T}$ and $\alpha = \mu_T\gamma_T = \mu_S(1+p_f m_{1K})\gamma_T \geq \mu_S\gamma_T$. Since the mean effective session time may change considerably. But with respect to (A14 b) and Fig. 11 we may conjecture that the assumption of $c_T^{-2} = \gamma_T \approx 1$ may support sufficient accurate interims estimations. Then (28) and (29) yield

$$\bar{F}_T(x,\gamma) = \frac{\Gamma[1,\mu_T x]}{\Gamma(1)} = \text{ and } m_{1I}(\Delta x) \cong \sum_{j=1}^{\infty}\int_{j\mu_T\Delta x}^{\infty} e^{-t}dt = \sum_{J=1}^{\infty} e^{j\mu_T\Delta x}.$$

Thus, by (29) and (A12c) we have

$$m_{1I}(\Delta x) \cong \frac{e^{-\mu_S(1+p_f m_{1K})\Delta x}}{1-e^{-\mu_S(1+p_f m_{1K})\Delta x}} \leq \frac{m_{1T}}{\Delta x} = \frac{1}{\mu_S(1+p_f m_{1K})}. \tag{A20}$$

References

[1] M. M. Zonoozi and P. Dassanayake, ‘User Mobility Modeling and Characterization of Mobility Patterns‘, IEEE Journal on Selected Areas in Communications, Vol. 15, No. 7, pp. 1239-1252, September 1997.

[2] D. Hong and S. S. Rappaport, ‘Traffic Model and Performance Analysis for Cellular Mobile Radio Telephone Systems with Prioritized and Nonprioritized Handoff Procedures’, IEEE Transactions on Vehicular Technology. Vol. VT-35., No. 3, pp.77-92, August 1986.

[3] W.-S. Soh and H. S. Kim,”QoS Provisioning in Cellular Networks Based on Mobility Prediction Techniques”, IEEE Communications Magazine, pp. 86-92, January 2003.

[4] Y. Fang and I. Clamtac, ‘Call Performance for a PCS Network‘ IEEE On Selected Areas in Communications, Vol. 15, No. 8, pp. 1568-1580, October 1997.

[5] Y. Fang, I. Clamtac and Y-B. Lin, ‘Modeling PCS Networks Under General Call Holding Time and Cell Residence Time Distributions’, IEEE/ACM Transactions on Networking, Vol. 5, No. 6, pp. 893-905, December 1997.

[6] Y. Fang, ‘Analytical Generalized Results for Handoff Probability in Wireless Networks’, IEEE Transaction on Communications, Vol. 50, No. 3, pp. 396-399, March 2002.

[7] Y. Fang, ‘Modeling and Performance Analysis for wireless mobile Networks: A new analytical Approach‘, IEEE/ACM Transactions on Networking, Vol. 13, No. 5, pp. 999-1002, October 2005.

[8] S. Zaghloul and A. Jukan, ‘On the Performance of the AAA Systems in 3G Cellular Networks‘ IEEE international Conference on Communications Glasgow, UK, 2007, pp. 2103 – 2108.

[9] S. Zaghloul and A. Jukan, ‘Modeling and Signaling Rate Estimation of the Diameter Protocol in All-IP Cellular Networks‘, Manuscript of an intended INFOCOM Contribution 2008, pp. 1-9.

[10] H. C. Tijms, ‘A First Course in Stochastic Models‘, John Wiley & Sons Ltd., 2003, 478 p.

[11] H. L. Hartmann, ‘The uniform estimation of the M/G/1 processor sharing response time distributions’ Eur. Trans. Telecomms. (ETT), 2008; 19: pp. 619-633.

[12] H. L. Hartmann, ‘Session Time Distributions of Cellular Networks with forced Terminations’, forthcoming Contribution under Progress.